DOCUMENT RESUME

ED 267 085 TM 860 160

AUTHOR Caudle, Sharon L.
TITLE Evaluation as a Process: The Formative-Summative Continuum.
PUB DATE Oct 85
NOTE 22p.; Paper presented at the joint meeting of the Canadian Evaluation Society, Evaluation Network, and Evaluation Research Society (Toronto, Canada, October 17-19, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Case Studies; Evaluation Criteria; Evaluation Methods; *Evaluation Needs; Evaluation Problems; Federal Programs; *Formative Evaluation; *Program Evaluation; *Research Methodology; Research Needs; *Summative Evaluation
IDENTIFIERS *Evaluation Research; *National Commodity Processing System

ABSTRACT

Rather than automatically presuming explicit conditions exist when designing an evaluation to fit the summative or formative mold, evaluators should think of an evaluation design as fitting between endpoints on an evaluation process continuum. Evaluators can blend techniques from both the formative and summative evaluation, matching actual program conditions, "sliding" along the continuum to meet a primarily formative or summative goal. Using a case study, this paper explores ways in which formative and summative techniques can serve such complementary roles in a single evaluation process. The evaluation of the National Commodity Processing System, a pilot program to reduce government food surplus stocks, illustrates how decision rules arising from presumed, but untested, summative conditions unnecessarily narrowed the research scope and the final product given to decision makers. This left little room for the depth and detail that qualitative data could offer. The continuum evaluation took a much broader approach to fit differing needs. Concepts, definitions, and models should not be ordered in advance, but selected to give balance to achieve the purpose at hand. Evaluating from the standpoint of a constantly creative continuum is one alternative approach to achieve that balance. (PN)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED267085

# EVALUATION AS A PROCESS: THE FORMATIVE-SUMMATIVE CONTINUUM

Sharon L. Caudle
Analyst, Office of Management and Budget
DPA candidate, The George Washington University

Paper presented at Evaluation '85, the international conference of the Canadian Evaluation Society, Evaluation Network, and Evaluation Research Society, Toronto, Canada, October 17-19, 1985.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. L Caudle
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 860 160

# EVALUATION AS A PROCESS: THE FORMATIVE-SUMMATIVE CONTINUUM

Sharon L. Caudle, D.P.A. candidate, George Washington University

Life is full of things we take for granted, that we assume as real without any evidence to establish their validity. In our thought processes, they are real--confirmed by previous fact patterns and similarities. They are the foundations for models that help in doing tasks and problem-solving. They become so engrained that we simply do not challenge their "reality" to us.

In conducting evaluations, we also take things for granted. Well-structured assumptions underlie our research, although the world we evaluate is not well-structured. We likely insist on applying evaluation typologies nicely categorized from idealized patterns of conditions--those conditions become reality for purposes of applying the typology. The danger is that we forget that the presumed typology realities are really just abstractions, not absolute properties.

This paper looks at two idealized typologies where we fully accept assumed disparate conditions. What begins as a black and white choice between summative and formative evaluation goals continues as a whole series of assumptions of reality, cast in terms of certainty and unambiguity. Once the formative or summative goal is known, then "summative" or "formative" condition assumptions start fencing-in the decision rule choices around research approach end points. These assumptions, normatively taken as true without proof, radically structure the evaluation approach along the disparate decision rule end points. Overall, the question is what the consequences are of dealing at the end points in evaluation.

Those assessing program results and impact think "summative" and restrict their approach to answer the question "Does the program work?" The evaluator's summative "reality mind-set" assumes a fully implemented, institutionalized program, where program goals are explicitly known and readily comparable against quantitative results. In lock-step, the evaluation design and production--the formulation of questions, the data collection, the analysis techniques, and the writing of the findings--have the stamp "made in the summative assumption mold" all over them.

At the other end of the evaluation typology spectrum, those who think "formative" take on a different set of evaluation "mind-set" assumptions. Digging into program management and improvement, the evaluator happily sets forth to look at implementation processes and modifications, asking "How can the intervention be better?" He or she assumes the intervention is still being implemented, and that some program components are likely not functioning as well as they should. The benchmark goals generally have been fleshed out, but the impact measures are still working themselves out.

It is doubtful we even state these assumptions up front as explicit evaluation decision rules--they are an invisible part of the evaluation's modus operandi. The invisible assumptions are further reinforced when we typically interchange typology descriptions: "impact" evaluations are synonomous with summative ones, and "process" evaluations are the same as formative.

"In the most concise terms, impact evaluation is the study of the extent to which goals have been achieved, whereas process evaluation is the study of how a program has been implemented" (Palumbo and Sharp, 1980:289). At the extremes, and in their conciseness, the impact/summative and process/formative invisible assumptions result in an unnecessary simplification, and at times incorrect evaluation approach. The typology abstraction obscures reality: we misread actual program conditions.

In this paper, a different method is explored. Rather than automatically presuming explicit conditions exist when we design an evaluation to fit the summative or formative mold, we should think of an evaluation design as fitting between endpoints on an evaluation process continuum. We can blend techniques from both the formative and summative sides of the street, matching actual program conditions, "sliding" along the continuum to meet a primarily formative or summative goal.

Using a case study, the paper explores ways in which formative and summative techniques can serve such complimentary roles in a single evaluation process. A food commodity demonstration project illustrates how decision rules arising from presumed, but untested, summative conditions unnecessarily narrowed the research scope and the final product given to decision makers.

## Defining the Continuum

Decision rules in designing an evaluation start at a very broad base. Edward Suchman (1967:39-41) discussed early on how evaluators should define the objectives of what is being evaluated in formulating evaluative research. The answers to certain questions would determine how the evaluation study was designed, for example, what is the nature of the content of the objective, e.g. changing knowledge, attitudes, and/or behaviors? Who is the target of the program? How is the objective to be attained?

Once the intervention's objectives are known under the Suchman model, we should understand how a program is supposed to work. Then decision rules organize the evaluation itself. They focus attention on data outcomes and uses (Ross, 1980:65-66). We explicitly take for granted certain rules within formative or summative ideal model frameworks which arise from assumptions of reality. These presumed conditions mandate decision rules on focus, research methodology, program stage, and program performance criteria. These rules illustrate constrained "either/or but not both" choices fitting the evaluation's overall impact or process scope. Figure 1 displays the major typology assumptions, and the resulting decisions as endpoints on a choice continuum.

Figure 1: Choices On the Continuum

| Assumption | Formative Endpoint | Continuum Choice | Summative Endpoint |
|---|---|---|---|
| FOCUS | Process ⟨----------⟩ Impact | | |
| | * input | * total production process | * output/ outcome |
| | * how can it be improved | * when does it work | * does it work |
| | * component assessment | * assess components within a whole | * holistic assessment |
| RESEARCH METHODOLOGY | Qualitative ⟨----------⟩ Quantitative | | |
| | | complimentary qualitative and quantitative measures | |
| PROGRAM STAGE | Implementing ⟨----------⟩ Implemented | | |
| | | tailor to program development stage | |
| PERFORMANCE CRITERIA | Forming ⟨----------⟩ Known | | |
| | * program component success | * components and overall success | * total effects against goals |

Focus Assumption: Impact Versus Process

The main distinction between formative and summative evaluation assumptions is what being studied. The assumed reality is that summative evaluations narrowly look at the end result; formative looks at the process that gets to that result. Two subassumptions further structure reality. The first is the question we need answered--does it work versus how can it be improved. The second sub-assumption which drives a wedge between the two typologies is the breadth of the assessment. Here it is presumed the formative evaluation assesses the functions of parts of a process, while the summative considers a whole intervention in almost a "black-box" structuring. Formative evaluations are free to monitor progress against objectives, and evaluate program design adequacy.

Research Methodology Assumption: Qualitative Versus Quantitative

It is assumed summative evaluations are highly likely to be quantitative in nature, using the "scientific" approach. The argument goes we must measure program effects quantitatively for credibility. Hard numbers demonstrate impact or lack of it. Frequently the data analysis is cumulative, presuming a door opening and shutting on the program under study. Formative evaluations, zeroing in on process as well as progress, can sample more of the qualitative spectrum.

They can experiment more with the quantitative part, too, taking parts of the program, or different time frames, mirroring program management concerns or changes.

Program Stage Assumption: Implementing Versus Implemented

It is assumed the new intervention is on its shakedown cruise when a formative evaluation is planned. The shakedown cruise is a continuation of the orginal policy formulation. The evaluator's role is re-tooling to ensure the program is implemented properly. Summative happens once a program is in place and the intervention is considered completely implemented and seaworthy. The role here is charting the cruise; the assumption is the program goals have not been modified, or modified to an extent unacceptable to decision-makers.

Performance Criteria Assumption: Known Versus Unknown Performance Measures

The presumed reality is summative evaluations stress goals, results, and effects--lining up measures of impacts against known benchmarks. While some may describe the strengths and weaknesses of the program, the performance measurement is not primarily intended for future program modification. Formative ones look at program procedures and modification. It is not critical, nor is it expected, that measures of impact be made. The interest is in how the program is being implemented, not necessarily how its impact measures up against a selected benchmark.

The point is the extremes in underlying typology assumptions place in our minds beliefs in reality which constrain the evaluation's resulting decision rules. However, between the extremes, the continuum approach permits great flexibility in the focus, research methodology, tailoring to the program stage of development, and performance criteria.

---

The argument for a continuum, or "blend" approach is not new. Suchman (1967:66) said "Making sense of the evaluative findings is the basic reason for adding a concern with process to the evaluation study. Otherwise one is left with the descriptive results of the evaluation, but without any explanation." Michael Quinn Patton (1982:4-48) suggests we should think of formative and summative models as ideas, not recipes or ideals. Joe Nay and Peg Kay (1982:95) say a formative evaluation system should develop as the intervention does. The system would 1) start monitoring as things begin to exist, 2) evaluate outcomes of processes as they occur, and 3) consider potential measures as impact models unfold.

This suggests a program never stops changing: a continous evaluation technique is needed to match the program changes. Ward Edwards and his associates (1983:145-146) caution us that change is continuous. First, the program's clientele and administrators change, reacting to program experience and their interaction. Second, the program itself will change. Third, the program's external environment will change, and lastly, program event and consequence knowledge accumulate.

These authors make the point a continuous evaluation technique is thus needed to match the program changes. Though a program's implementation may fit in an "institutionalized" category, no program is entirely stable across all

assumptions. There are always changes, large and small, at the margins.

This perspective is perhaps even more true in the case of demonstration projects. As their name implies, they are intended to demonstrate the value of some intervention. However, they are even more prone to the impacts of the change process. A program is not set at one point in time in these cases: it is constantly evolving as it settles into a policy and procedural routine defined by the implementation process. A strictly summative approach likely will not fit well, even at the point the case is made that the demonstration project is fully implemented--"institutionalized."

For well-established, ongoing programs, summative evaluations assume goal achievement--the program has delivered whatever it was supposed to to the appropriate targets. However, even ongoing program monitoring serves a formative process as it provides information on how well the program is meeting its intervention design. The critical "need to know" information to monitor effectiveness is how adequately implemented the intervention is. The evaluation design should comprise both formative and summative approaches to nail down both an acceptable intervention process at work, and outcome results.

## The Case Study: A Program Explanation

In June 1983, USDA's Food and Nutrition Service (FNS) unveiled a new pilot program called the National Commodity Processing System (NCP) to help reduce massive surplus stocks of government-owned dairy and other agricultural products. Under the pilot program, FNS contracted directly with food companies (called processors) to convert the surplus commodities into end products like pizza and puddings. The end products are sold at lower prices to eligible recipient agencies (RAs) such as schools and charitable institutions, reduced for the commodity value portion of the products.

NCP was a parallel program of long-standing State-run processing programs which did not entirely cover all the processors or recipient agencies. NCP was to fill the market gap. Some States were reluctant to expand their list of processors or even enter into processing agreements at all if they had not done so already. Some non-participating State distributing agencies did not have staff to negotiate the agreements and ensure conformity with the USDA monitoring requirements. Other State agencies limited the recipient agencies participating in processing programs, further restricting sales.

USDA decision makers expected the following benefits from NCP:

o Provide the opportunity for many recipient agencies to receive processed end products. This would be particularly true in those States where the distributing agency refused to enter into processing agreements or where the State agency only permitted schools to participate in processing. Charitable institutions, Nutrition Programs for the Elderly, and other outlets eligible for food distribution could participate, and reduce their food costs.

o Provide a reduction in the paperwork for both processors and recipients associated with entering into many State contracts. Eliminating duplication and overlap of effort could result in significant cost reductions. It was anticipated that if there was a single entity processors could deal with,

they would be more prepared to enter into processing agreements.

o Potentially remove an additional 100 million pounds of dairy products held in CCC storage. This would greatly reduce the high cost of storage. Since these commodities go out of condition within one to two years, any measure to attempt to bring supply in line with demand was critical.

o Provide an incentive for processors to develop new products for use by eligible recipients which would, in turn, lead to increased demand for surplus commodities.

In sum, the program logic was quite straightforward:

| | |
|---|---|
| Program Operation | * recipient agency registration for NCP participation |
| ↓ | * State agency assistance/dissemination of NCP information |
| ↓ | * processor service delivery system |
| ↓ | * end product development |
| ↓ | * use of recipient agency listings for market development |
| ↓ | * commodity ordering and end product distribution |
| ↓ | * FNS monitoring and evaluation through automated management information system |
| ▽ | |
| Measureable Outputs | * number/type of processors approved |
| | * number/amount food orders |
| ↓ | * number/amount food sales |
| ▽ | * number/types RA's registered |
| Intermediate Results | * expanded RA's participate in processing |
| | * additional processors in processing |
| ↓ | * innovative end product development |
| ↓ | * program networking |
| ▽ | * supplement to State processing |
| Direct Effects | * primary: net increase in bonus dairy commodities moved out of storage |
| | * secondary: availability of cost savings to maximum eligible recipient agencies |

The logic relied on complex program implementation and coordination on both intergovernmental and public/private levels. Very specific role and performance were laid out: State agencies were the vital link in recipient agency knowledge of, and participation in, the NCP System. Processors were to venture into new markets with new products, essentially responsible for the success of the program since they were the linchpin to the independent service delivery system. FNS was to serve as the facilitator of this process, ensuring the processor had recipient agency listings to market with, and ensuring the flow of commodities for production as the processor made sales.

## Contrasting Assumptions: Summative and Continuum Approaches

The June interim rules authorized the program for a two year period. In late 1984 and early 1985, FNS' Office of Analysis and Evaluation (OAE) conducted an evaluation of the program and recommended it end June 30, 1985. The OAE

evaluation followed a classical summative approach (USDA, 1985). The evaluators assumed explicit conditions in formulating their decision rules, noted below under SUMMATIVE. This is contrasted with CONTINUUM ones.

The continuum evaluation used information that replicated the sources and timing of data available to the FNS evaluators, supplemented by my personal observations and records as the project manager from October 1983 to April 1985. NCP performance data included recipient agency registration, commodity sales and orders, processor agreement records, and other briefing material as of April 9, 1985.

Focus:

SUMMATIVE:
The decision-makers' charge was seen as obtaining "go, no go" advice on continuing the pilot program, not looking at the development of the pilot program to its current stage. The focus was impact assessment, not process.

CONTINUUM:
As the pilot period was drawing to a close, program managers anticipated there would be political pressure to continue NCP either as a pilot program or as a permanent one. The focus was expanded to include the question "when does it work" so program modification and process issues could be covered. Instead of looking at the program holistically as the summative one would, the continuum approach could assess program components as well as overall program impacts.

Research Methodology:

SUMMATIVE:
Mostly quantitative data would comprise the research methodology. Internal FNS statistical records would be used as almost the sole source of information. The preliminary research design included a survey of State agencies on both formative and summative characteristics. However, the summative evaluation used only specified State agency quantitative data, and did not include audit and operational data that was primarily formative in nature. The lead evaluator believed the survey data was too qualitative in nAture--it did not provide statistics, only anecdotes. He also questioned the utility of many of the questions which asked for State opinions on possible changes to the program. His view was that "the real story was what did NCP do, the impact on State programs. What to do with NCP is a side issue" (Goodwin, 1985).

CONTINUUM:
Qualitative data and other descriptive data was desired to supplement the available quantitative data. The State agency survey data covered process and impact issues and thus was analyzed for inclusion in the results. Descriptive data from program managers, State agencies, and processors supplemented quantitative data.

Program Stage:

SUMMATIVE:
It was assumed the program was fully implemented as designed and thus any program non-performance would be charged to program failure.

CONTINUUM:

The evaluation would be tailored to the program development stage. The degree of implementation would share coverage with program theory concerns.

Performance Criteria:

SUMMATIVE:
Explicitly stated regulatory program outcome goals would be solely compared against known cumulative results. Expected outcome criteria were known. Any changed goals or different program effectiveness measures that might have derived from pilot program operations would not be explored. The June 1985 FNS evaluation stated the NCP goals as 1) increasing the consumption of surplus dairy products by attracting major food companies not participating in State processing, serving outlets and States previously unserved, and promoting innovative dairy commodity products and 2) supplementing existing State processing programs.

The impact measures would be in sales against goals, the extension to non-traditional recipient agencies, and the supplementation of State bonus processing programs. Data on State processing of donated commodities for 1982-83 became the baseline data for comparison against the 1983 implemented NCP program. To supplement limited availability of NCP sales data, the evaluators used recipient agency registration and processor food orders.

CONTINUUM:
The regulatory program logic would be compared to actual operations, and results compared to regulatory goals. Since the degree of implementation would be an issue, cumulative results against a baseline measure would be approached cautiously. Both program implementation and program theory would be under study. Since the program logic indicated recipient agency savings was a performance measure, that would be examined.

The SUMMATIVE decision rules polarized around the summative endpoint; the CONTINUUM spread across the decision rule band. Figure 2 compares the outcome of the different approaches of the summative and continuum evaluations.

Figure 2: Comparison of Continuum and Summative Assumptions

| Assumption | Summative | Continuum |
|---|---|---|
| FOCUS | impact assessment | impact and process |
| RESEARCH METHODOLOGY | quantitative: internal records | quantitative/qualitative: internal records, external sources, managers' views |
| PROGRAM STAGE | fully implemented | implementation not assumed |
| PERFORMANCE CRITERIA | regulatory-stated goals against cumulative results; program theory examined; performance criteria known | regulatory goals and goals deriving from program operations; program theory and implementation examined; performance criteria not assumed |

## Contrasting the Summative and Continuum Results

The summative evaluation report was highly quantitative and focused on the bottom line of program impact. In contrast, the continuum results provided a broader picture of program operations and impact. By looking at the conditions under which NCP did, and did not work, at its components as well as the whole program, using "numbers" as well as descriptive data, considering the program development stage, and considering performance measurement which considered all of these factors, the continuum approach provided additional perspectives for the decision-makers. The results of each are contrasted below.

### Program Implementation and Program Theory

SUMMATIVE:
NCP was fully operational during the pilot period and met all of its program design conditions. Its failure would be the result of program theory problems.

CONTINUUM:
In the first two years, program managers described a myriad of program development and implementation problems and unanticipated results. Poor initial program planning failed to anticipate the resources needed to administer a national program serving tens of thousands of recipient agencies:

- o Implementation fast tracking of created serious problems with program accountability and the automated national management information system.

- o Volume processing and profit margins required that large processors, normally those attracted to NCP, concentrate on school sales. Small recipient agencies, such as child care programs, were hard to market and distribute to compared to the large school ready-made market. Many processors found marketing to recipient agencies other than the traditional school market had to wait until the school market base was secured. Thus relying on processors to "reach out" to all eligible recipient agencies as the program logic required was a problem.

- o The NCP program logic assumed fairly quick innovative end product development. However, end product development took from six months to a year in most cases. Processors normally would set up a traditional food line with the commodities such as pizza, creating a track record of performance with the recipient agencies and a foundation for selling the new products.

- o Those processors entering NCP for the first time were frequently subsidiaries of larger companies who, with their large organizational structures, needed a substantial lead time for decision-making. NCP, being a pilot program, was an unknown quantity for new companies; those with experience in State processing were in a much better position to make the transition to NCP, or supplement current State processing activities.

- o The program logic assumed that recipient agencies would be registered for processors to begin their market development. In fact, the regulations required proof of marketability to these recipient agencies before commodity food orders were placed. However, many State agencies did not

distribute NCP registration cards to their recipient agencies. Those that did distribute the cards were not provided program fact sheets which would explain the program to potential registrants. Ironically, the registration card itself did little to explain to recipient agencies how the NCP program operated.

FNS later developed an explanatory letter for State agency use and periodically sent approved processor lists to FNS regional offices for distribution. FNS also sent an informational letter to newly registered recipient agencies. However, lists of processors were not provided directly to recipient agencies--the potential market. In the last few months of the pilot period, a mailing to over 23,000 recipient agencies offered to send a processor listing to any recipient agency requesting them. Much to the program managers' surprise, over 10,000 requests were received. Awaiting a decision on termination of the program, FNS senior officials decided to hold off on the processor listing mailing indefinitely.

o FNS substituted a processor's State processing performance and/or a discription of potential markets for the rule required proof of marketability. Inadvertently, this process may have contributed to encouraging processors active in State processing to move to or supplement the State activity with NCP.

o A running confrontation with many State program officials emerged. These State distributing agencies, running their own processing programs with minimal USDA oversight, saw NCP as carpet-bagging into State territory. Their power and control over recipient agencies eroded as FNS or the processors, through the independent delivery system, dealt directly with local recipient agencies. There also was no way to identify and prevent processors from reporting duplicate sales under both NCP and State programs.

o FNS regional and State staff were left out of the program planning, despite the need for their assistance in actual operations once the program got started.

NCP Sales

SUMMATIVE:
NCP sales during the two year pilot period will be much lower than the anticipated level of 100 million pounds. Sales during 1983-1984 totaled only 6.8 million pounds. Food orders rose from 9 million pounds for all of 1983-84 to 27 million in the first seven months of 1984-85. However, one-half of these orders are the result of displaced State processing.

CONTINUUM:
A caveat in drawing conclusions from the sales data is the problem of updating the sales masterfile. Processors report sales monthly, but can cover updates of sales back to the beginning of the agreement period. Delayed reporting, as well as delays in entering the data into the NCP management information system provide flawed "snap-shots" at any point in time. Particularly with voluminous monthly reports, program staff noted that delays in entering the data into the system could extend several months.

In the first first year of NCP operation, recipient agencies purchased NCP

products containing 7.093 million pounds of bonus processing commodities. About 84 percent of this total was purchased in the last six months, reflecting both start-up delays and the normal purchase cycle of the largest recipient agency group--schools. Thirteen percent was purchased in January, 15.1 percent in February, 14.8 percent in March, 15.7 percent in April, 15.3 percent in May, and 9.5 percent in June. Despite the delay in sales reporting and data entry, 8.9 million pounds were sold in the July 1984 through February 1985 period as of the sales reported on April 1985 FNS reports. Of this total, "closed out" reporting months comprise 62.6 percent--22.6 percent was sold in September, 23.8 percent in October, and 16.2 percent in November. However, a goal of 100 million pounds movement was not achievable by the end of the pilot period.

Cost Savings

SUMMATIVE:
The Federal cost of administering NCP considerable exceeds savings due to reduced food storage. Annual storage savings are $276,767; the cost of running the program is approximately $690,000.

CONTINNUM:
In addition to reduced food storage costs, the program logic included the impact of cost savings to recipient agencies. Recipient agency benefits were considerable. Using the first 6 months of second year NCP sales and June 1984 commodity prices, the annual projected savings to recipient agencies was $16.7 million. Using commodity deliveries for the first 8 months as a projection of anticipated sales, the annual savings rise to $40.4 million.

Recipient Agency Registration and Sales

SUMMATIVE:
A large portion of NCP commodities are purchased by a small number of schools located in a handful of States. Only 7 percent of the non-school recipient agencies made NCP purchases, one percent of all active recipient agencies bought 31 percent of the commodities, and 5 States (Illinois, Indiana, Missouri, and Ohio) account for 70 percent of NCP sales.

NCP recipient agency registration is concentrated geographically and by recipient agency type. Five states accounted for 55 percent of the registered recipient agencies. Schools are 60 percent of the agencies. Registration in the second part of the pilot period, an indicator of future sales, grew slowly.

CONTINUUM:
Of the potential universe of 55,000 recipient agencies, a little under one half were registered for NCP participation. School food authorites composed 34 percent of the total number registered, followed by charitable institutions, 23 percent; independent public schools, 17 percent; private schools, 9 percent, child care feeding programs, 6 percent, elderly nutrition programs and correctional facilities tied at 4 percent, and summer food service/camps, 3 percent.

School food authorities composed 53 percent of the total sales activity, followed by independent public schools, 24 percent; private schools, 10 percent, charitable institutions, 7 percent; correctional facilities and child care feeding programs tied at 2 percent; and elderly nutrition programs, 1 percent. It is not surprising that school food authorities and independent schools

comprise the bulk of those active. These recipient agencies are the largest immediately profitable markets for processors, representing thousands of meals to school children.

For comparison, the continuum study used as a baseline April 1984 data and compared it to April 1985. The April 1984 figures represent at least six months of program operations following program start-up in mid-summer 1983. Looking at the detail of those registered the study first examined the expansion of processing to non-traditional State processing recipient agencies. The study measured this in terms of school food authorities versus non-school food authority registration.

Nationally, the ratio of school food authorities to non-school food authorities actually increased slightly from .51 to .53. However, fifty-five percent of all State agencies increased their ratio; 41.1 percent decreased; and 3.9 percent stayed the same. However, this information cannot be put in the context of program measurement criteria. The proper ratio never was specified in the program development, nor was the universe by State of school food authorities and non-school food authorities ever tabulated.

There was a 40.3 percent increase in registrants from the April 1984 period to the April 1985 period and a 427.5 percent increase in active registrants. As of April 1984, 8 percent of the registered recipient agencies were active. As of April 1985, about 30 percent of registered recipient agencies were active. Altough decision-makers did not specify what rate of growth or actual registration was reasonably expected at various points in the development of the program, the forty percent increase appears to be substantial. In the two year period, less than ten recipient agencies asked to be removed from the registered list.

Figure 3 shows April 1984 and April 1985 registration and sales activity as a percentage of national registration and active sales for each FNS Region (NERO-Northeast; MARO-MidAtlantic; SERO-Southeast; SWRO-Southwest; MPRO-Mountain Plains; WRO-Western).

Figure 3. Regional Registration and Sales Activity 4/84 and 4/85

| | April 1984 | | April 1985 | |
|---|---|---|---|---|
| Region | % Nat Regist | % Nat Sales | % Nat Regist | % Nat Sales |
| NERO | 12.0 | 4.4 | 10.6 | 4.0 |
| MARO | 14.0 | 6.0 | 11.6 | 6.4 |
| SERO | 6.0 | 0.8 | 10.4 | 11.4 |
| MWRO | 31.5 | 69.1 | 30.4 | 51.4 |
| SWRO | 5.8 | 0.6 | 9.6 | 7.1 |
| MPRO | 18.0 | 11.3 | 16.5 | 13.8 |
| WRO | 12.8 | 7.7 | 10.9 | 6.1 |
| NATIONAL | 100.0 | 100.0 | 100.0 | 100.0 |

Thus there were swings in SERO and SWRO to a larger piece of the sales pie. To further refine the comparison given the delay in program start-up, the months of October, November, and December 1983 were used as a baseline for comparison with the same months in 1984 for the second year of the pilot's operation, though the

second year's data may still be understated because of reporting delays, but it gives us a feel for specific State changes. Figure 4 shows the top States for these months.

Figure 4 Top Sales States: October-December 83; October-December 84

| State (dropped bonus) | | 10/83-12/83 Period % Nat Sales | Rank | 10/84-12/84 Period % Nat Sales | Rank |
|---|---|---|---|---|---|
| Illinois | (July 1983) | 39.8 | 1 | 11.6 | 2 |
| Michigan | | 22.0 | 2 | 9.1 | 4 |
| Indiana | (July 1983) | 8.4 | 3 | 4.6 | 6 |
| Missouri | (July 1984) | 3.6 | 4 (tie) | 3.2 | 8 |
| Arizona | (July 1984) | 3.6 | 4 (tie) | 2.6 | 9 |
| Washington | (did not have SP) | 3.6 | 4 (tie) | 1.2 | 10 |
| Ohio | | 1.7 | 5 | 14.1 | 1 |
| Texas | (July 1984) | .4 | 6 (tie) | 10.1 | 3 |
| Alabama | (July 1984) | .4 | 6 (tie) | 4.2 | 8 |
| Tennessee | (July 1984) | .3 | 7 | 4.4 | 7 (tie) |
| Florida | | .0 | 8 (tie) | 9.0 | 5 |
| Minnesota | (July 1984) | .0 | 8 (tie) | 4.4 | 7 (tie) |
| | | 83.8 | | 78.5 | |

As is apparent from Figure 4, there is considerable change in results in the second year, but NCP sales by and large are still made to States not involved in bonus processing for these comparison periods. The MWRO represents the lion's share of registrants and sales activity, but is apparently losing ground in these "snap-shots" of program performance.

Processor Participation

SUMMATIVE:
NCP sales and food orders are concentrated among a few processors. Five companies out of the 55 approved to participate in 1983-84 made 87 percent of all sales; 10 made 97 percent. Ninety percent of the food orders in both years of the pilot period came from the same ten companies.

CONTINUUM:
The program logic relied on an extensive processor distribution system, thus increasing the number of recipient agencies receiving the benefits of commodity assistance through processing agreements. There was no restriction on the type or number of processors; the rules specifically noted that any food processor who could manufacture commodities into a finished product was eligible to apply. In fact, FNS operational staff used the applicant processor's State processing performance as the key measure of anticipated NCP performance. Processors already in State processing for the most part had already established product lines, and had production plants capable of handling the commodities. As was mentioned earlier, companies which had not participated in any commodity processing needed substantial research and development time.

In the first year of the pilot program, 65 processors applied and 55 were approved. In the second year, 82 had applied as of April 1985, and 64 of these had been approved. The large majority of the applicants for both periods had or

were currently participating in State processing. The largest concentration of processors were in the Midwest and Mid-Atlantic regions, in traditional food production and distribution areas.

Six withdrew in the second year. The main reasons for withdrawing were problems in developing new products and distributing them, the uncertainty of NCP continuation, and pressure from States to have either NCP or State processing, but not simultaneous agreements. In an October 1984 report which looked at 65 of the 69 processors approved or awaiting approval, many processors reported problems resulting from NCP-State processing co-existence. Nineteen indicated they were prevented by 25 States from entering into or renewing State agreements because of NCP participation. Five States refused to allow a processor to have both NCP and state processing contracts in their State. Some processors were affected by only 1 of the 25 States; one processor was affected by 16 States.

Seven processors had 91.2 percent of the national sales in the first year of NCP operations. These same seven had 87.9 percent of the sales for the first six months of the second year. They also accounted for about the same percentages of all deliveries. Using FNS records by processor, total processor sales increased about 70 percent from the first year to the second year, projecting the first six months of the second year sales to an annual number. The sales pound comparisons were 7.1 million pounds in the first year to a projected 12.0 million pounds in the second year. Deliveries were projected to total over 29 million pounds for the second year.

State Processing Impact

SUMMATIVE:
Thirty-two States operated bonus processing programs in 1984-85, a net drop of six since the inception of NCP. Eleven States completely dropped bonus processing, citing NCP overlap as one of the reasons; the dropping States were concentrated regionally with 4 in the Midwest and 3 in the Southeast. Five other states picked up bonus processing; these were scattered across the nation.

NCP caused some States to curtail processing activities for reasons such as processors participating in NCP only, or States not approving NCP processors. NCP's overall effect was to displace State processing. However, for individual States the effects vary. In different States, it can be demonstrated that NCP supplemented State processing, displaced State processing, or resulted in a net decrease in combined State-NCP processing. Of the 5 States that account for 70 percent of all NCP sales, 3 have completely dropped bonus State processing, and the other two have decreased their number of processing contracts by nearly half.

CONTINUUM:

A critical indicator of NCP impact is the displacement of State processing, noted in the previous section on recipient agency registration. The FNS Fall 1984 NCP Evaluation Survey gathered data on State processing agreement changes, both for entitlement commodities such as meat products and vegetables, and the bonus processing. Fifty-eight State agencies provided information; 72.4 percent of these in the period July 1982-June 1983 had processor agreement activities, 78.9 percent in July 1983-June 1984, and 69 percent in the last pilot year of NCP operations. The other respondents were administering agencies for the most part who might register recipient agencies for participation in the State's

processing program, run by another State organization.

Figure 5 shows by region the number of State agencies with agreements for the comparison periods of before NCP (July 1982-June 1983), the first year of NCP (July 1983-June 1984), and the last year of NCP operations up to March 1985. The number of actual agreements, both bonus and entitlement, are shown in parentheses.

Figure 5 Number of State Agencies Participating in Processing by Region
(Number of total agreements in parentheses)

| | 7/82-6/83 | | | 7/83-6/84 | | | 7/84-3/85 | | |
|---|---|---|---|---|---|---|---|---|---|
| Region | Bonus | Entit | None | Bonus | Entit | None | Bonus | Entit | None |
| NERO | 6 | 6 (237) | 0 | 6 | 6 (250) | 0 | 6 | 6 (253) | 0 |
| MARO | 4 | 5 (264) | 4 | 4 | 5 (315) | 4 | 5 | 5 (262) | 4 |
| SERO | 8 | 7 (178) | 0 | 8 | 7 (184) | 0 | 5 | 6 (120) | 2 |
| MWRO | 6 | 6 (86) | 0 | 4 | 5 (73) | 1 | 2 | 3 (28) | 3 |
| SWRO | 4 | 3 (63) | 1 | 5 | 5 (52) | 0 | 3 | 3 (16) | 2 |
| MPRO | 6 | 7 (47) | 3 | 8 | 7 (65) | 1 | 7 | 8 (32) | 1 |
| WRO | 4 | 3 (29) | 4 | 4 | 3 (50) | 4 | 4 | 5 (85) | 2 |
| NATIONAL | 38 | 37 (904) | 12 | 39 | 38 (995) | 10 | 32 | 36 (796) | 14 |

Of the 53 State distributing agencies responding, 71.7 percent had bonus agreements, 69.8 percent had entitlement agreements, and 22.6 percent did not have any for 7/82-6/83, the baseline period. In the first year of NCP operations of 7/83-6/84, the percentages had changed to 73.6, 71.7, and 18.9 percent respectively. In the last agreement period, the percentages were 60.4 percent having bonus agreements, 67.9 percent having entitlements, and 26.4 percent with no processing agreements.

In terms of number of agreements, there was a jump from 904 to 995 in the first year of NCP operations, then a radical drop to 796 as a number of States dropped agreements, either totally or partially. Just looking at the time periods, there is a likely linkage between NCP and States dropping or curtailing processing agreements, indicated by the obvious slippage in agreements in areas where States no longer were involved in bonus commodities.

The Survey asked respondents if a decision had been made to eliminate or curtail and Statewide processing involving bonus dairy commodities for reasons other than for nonperformance or noncompliance in each of the three periods. Six percent (N=3) of all respondents had made such a decision in the baseline period, 15.4 percent (N=8) in the first year of NCP operations, and 48.1 percent (N=26) in the second year of operations. Of those with 1984 agreements, 7.5 percent (N=3) did so in the baseline period, 15 percent (N=6) in the first NCP

year, and 48.7 percent (N=19) in the second year. When asked about the reasons for curtailing or eliminating in any or all of the three years, the respondents noted several reasons. Figure 6 shows the percent of those answering "yes" to the curtailing/eliminating question. (TFAP is the Temporary Emergency Food Assistance Program implemented at the same time as NCP.)

Figure 6 Reasons for Curtailing/Eliminating State Bonus Processing

| Reason | All Survey Respondents | With 84 Agreements |
|---|---|---|
| increased TEFAP workload | 18.5 | 15.0 |
| NCP duplicated services | 70.4 | 60.0 |
| NCP put SP at compet. disadvant | 22.2 | 15.0 |
| lack of staff | 33.3 | 30.0 |
| lack of federal technical assist | 7.4 | 10.0 |
| processors did not renew | 22.2 | 30.0 |
| States dropped NCP firms | 29.6 | 40.0 |
| NCP value pass thru not OK | 29.6 | 25.0 |
| RA confusion--2 programs | 33.3 | 15.0 |
| loss SA control over RAs | 14.8 | 10.0 |
| other reasons | 3.7 | 5.0 |

Although the survey did not ask for reasons by period, certainly the references to TFAP and NCP can be limited to beginning with the June 1983 timeframe. By far, NCP duplication of services is the main reason cited for the State curtailment/elimination. NCP-related problems account for a large percentage of other reasons.

Recommendations

SUMMATIVE:
The evaluation briefly highlighted four policy options for decision-makers. One, NCP could be extended either temporarily or permanently with continued voluntary State processing. Second, FNS could return to voluntary State processing with enhanced FNS technical assistance functions. Third, FNS could mandate State processing or unrestricted recipient agency participation. Lastly, FNS could develop a new system under which FNS and States would share responsibilities. A short discussion of the advantages and disadvantages of these options concluded the evaluation.

CONTINUUM:
The FNS evaluators chose not to use the data provided by State agency survey respondents on program redesign. However, many of the areas States mentioned were direct comments on problems in program implementation and program theory. As such, they richly supplement the cumulative quantitative data the FNS evaluation heavily relied on. This material could have been valuable to decision-makers in the discussion on policy options.

Respondents were asked their preferences on specific alternative processing programs, assuming there was not both State and NCP agreements. Twelve percent of all respondents believed FNS should return to State processing as the sole processing program. A little over 3 percent thought that FNS should institute a Federal requirement that States conduct a processing program, replacing the current voluntary program. Thirty-one percent believed that FNS should keep the

State and NCP agreements as is. A little less than 28 percent believed that FNS should establish a national inventory system to support State processing. States would approve processor agreements and monitor performance while FNS would order food and maintain inventories for each processor. Seventeen percent believed FNS should develop a "blend" program of State and NCP processing with clear Federal and State responsibilities. Nine percent indicated no preference.

Those respondents which were States with 1984 processing agreements had similar percentages. Eighteen percent believed State processing should be the sole program; 5 percent, mandated State processing; 22.5 percent, State and NCP as is; 32.5 percent, a national inventory supporting State processing; 20 percent, a "blend" program; and 2.5 percent, no preference. The indication from these comments is that State agencies prefer some program with a national involvement, but not a return to the status quo of voluntary State processing. The implication might be that despite the problems of the NCP program, States do see the value of a national component in processing.

Respondents were also given a list of possible objectives crucial for any processing program and asked to rank them from 1 to 5, with 1 being most important to 5 being least important. Figure 7 shows the ranking.

Figure 7 Ranking of Crucial Processing Objectives
(median: 1 most important; 5 least important)

| Objectives | All Respondents | With 84 Agreements |
|---|---|---|
| donated food accountability | 1.6 | 2.3 |
| reduced cost to recipient agencies | 1.7 | 1.7 |
| maximize donated commodity use | 1.8 | 1.8 |
| effective claims recovery | 2.1 | 2.2 |
| recipient agency service | 2.2 | 2.1 |
| minimal administrative burden | 2.5 | 2.0 |
| available to all recipient agencies | 2.6 | 2.5 |
| minimal ineligible sales | 2.7 | 2.7 |
| intergovernmental coordination | 2.7 | 2.8 |
| product variety | 3.3 | 3.7 |
| private sector participation | 3.3 | 4.5 |
| product innovation | 3.5 | 3.4 |

What is of interest in this ranking is how the NCP program objectives of product innovation, product variety, and private sector participation basically are ranked of little importance compared to areas of administrative accountability and control. Also, recipient agency service and cost savings receive a relatively top ranking whereas the summative evaluation discounted this as a program objective. The message in these State rankings is that NCP may have objectives which run counter to the preferred objectives of those who are involved in State processing.

The Survey asked respondents to indicate their assessment of the expected outcomes of several processing program options. State processing and NCP as is received the highest percentage seeing it as reducing CCC inventories, serving the widest range of recipient agencies, promoting product innovation, minimizing administrative burden, providing the widest range of products for recipient

agencies, and encouraging private industry participation. Voluntary State processing without NCP, a return to the status quo option was viewed by the highest percentage as maximizing the accountability for donated foods, minimizing sales to ineligible recipient agencies, and minimizing intergovernmental overlap.

The continuum evaluation has strongly emphasized that NCP implementation in terms of those program design laid out in the regulations was very questionable. However, even if this flawed implementation was satisfactory to decision-makers, it is likely the results still would be a "fatal remedy" (Sieber, 1981). There was no mechanism to prevent State agencies from dropping their voluntary programs. Indeed, systemic pressures to reduce recipient agency confusion, reduce State costs or staff burdens and move them to higher priority programs, and meet the processors need for a standardized program inexorably moved NCP to the policy disaster result. It also appeared to have objectives strongly at odds with those that the Survey respondents viewed as critical in any processing program. Without a more overt recognition of those objectives, NCP in any form may be viewed as a "valueless" program.

Thus, NCP as a separate program does not appear to be a viable option unless restrictions are placed on State processing to prevent the State displacement. A stronger course would be to strengthen State processing.

---

## Discussion of Contrasts

Figure 8 summarizes the results of the summative and continuum evaluations on various components.

Figure 8: Summary of Summative and Continuum Results

| Evaluation Components | SUMMATIVE | CONTINUUM |
|---|---|---|
| program implementation | o complete/adequate | o incomplete/questionable degree of adequacy |
| performance objectives | o known | o partially known |
| source of goals | o regulations | o regulations/program operations |
| performance baseline measures | o known | o unknown in most cases |
| acceptable data | o quantitative | o qualitative/quantitative |
| period of performance comparisons | o cumulative | o "closed out" periods in agreement years |
| focus | o outcomes | o process and outcomes |
| program theory | o flawed | o flawed |
| recommendation | o not selected | o selected |

This paper began with the premise that any evaluation should question its underlying assumptions, and draw from a pool of techniques and methodologies across a wide spectrum. The formative or summative purpose should not restrict the pool to narrow choices because of an ideal model. The summative evaluation was the result of that narrowing of choices. The evaluation decision rules left little room for the depth and detail that qualitative data could offer from the

State survey, nor could it open the evaluation to comparisons within agreement periods or questioning NCP's implementation. The continuum evaluation, on the other hand, took a much broader approach to fit differing needs.

More importantly, the summative conclusions in some areas were questionable as the continuum results demonstrate. This illustrates the very specific consequences which can result from unquestioning acceptance of assumptions at the formative-summative endpoints. Those consequences are shown in Figure 9:

Figure 9: Consequences of "Endpoint" Evaluation Assumptions

| Assumption | Formative Endpoint | Summative Endpoint | Consequences |
|---|---|---|---|
| FOCUS | Process | <---> Impact | Misinterpretation of results. |
| RESEARCH METHODOLOGY | Qualitative | <-> Quantitative | Ignore valuable data. |
| PROGRAM STAGE | Implementing | <-> Implemented | Incorrect comparison bases. |
| PERFORMANCE CRITERIA | Forming | <----> Known | Mask internal and external impacts. |

The summative evaluation provided only pieces of a complicated event. Meaningful, valid data on process and outcomes slipped like quicksilver through the summative decision rules' hands. The evaluation had to be as "black and white" as possible so decision makers were faced with a clear choice. However, the continuum information would have been extremely valuable in giving policy options to agency decision makers, and members of Congress who got involved in legislating an extension of the program.

NCP did not need to be maintained for its program objectives to be met. The summative evaluation left FNS decision-makers without any contingency plan support. Continuum evaluation provides support to improve programs if the political winds are such demonstration or apparently "sunseted" programs linger on. Quite simply, decision makers can use outcome evaluations to improve practices, but only if the research is designed for that possible end.

Michael Quinn Patton (1982:52-53) says

> [P]ractice is inevitably undergirded by conceptual distinctions that are far from esoteric. Whether one is aware of it or not, practice is guided by, indeed, controlled by--the concepts, definitions, and models that order that world upon which one is "practicing." To ignore the conceptual distinctions that direct the day-to-day decisions taken in the [evaluation] field is to be controlled and limited by those conceptual predispositions.

As evaluators, we need the concepts, definitions, and models to be free-flowing. They should not be ordered in advance, but selected to give us balance in the order we need in achieving whatever the purpose at hand. Evaluating from the standpoint of a constantly creative continuum is one alternative approach to achieve that fine balance.

## Selected References


Ward Edwards, Marcia Guttentag, Kurt Snapper, "A Decision-Theoretic Approach to Evaluation Research," in Handbook of Evaluation Research, Elmer L. Struening and Marilyn B. Brewer, Society for the Psychological Study of Social Issues, Sage Publications, Beverly Hills, 1983, 139-184.

Edie N. Goldenberg, "The Three Faces of Evaluation," Journal of Policy Analysis and Management, Vol. 2, No. 4, 515-525.

David Goodwin, Food and Nutrition Service, interview by Sharon Caudle, January 5, 1985.

Joe Nay and Peg Kay, Government Oversight and Evaluability Assessment, Lexington Books, 1982.

Dennis Palumbo and Elaine Sharp, "Process Versus Impact Evaluation of Community Corrections," The Practice of Policy Evaluation, ed. David Nachmias, St. Martin's Press, New York, 1980.

Micheal Quinn Patton, Qualitative Evaluation Methods, Sage Publications, Beverly Hills, 1980.

_____, Practical Evaluation Methods, Sage Publications, Beverly Hills, 1982.

_____, Utilization-Focused Evaluation, Sage Publications, Beverly Hills, 1978.

John A. Ross, "Decision Rules in Program Evaluation," Evaluation Review, Vol. 4 No. 1, 59-74.

Paul Sabatier and Daniel Mazmanian, "The Conditions of Effective Implementation: A Guide to Accomplishing Policy Objectives," Policy Analysis, Vol. 5, No. 4, Fall 1979, 491-504.

Edward A. Suchman, Evaluation Research, Russell Sage Foundation, New York, 1967.

Mary Ann Scheirer and Eva Lantos Rezmovic, "Measuring the Degree of Program Implementation: A Methodological Review," Evaluation Review, Vol. 7, No. 4, 515-528.

Sam D. Sieber, Fatal Remedies: The Ironies of Social Intervention, Plenum Press, New York, 1981.

USDA, Food and Nutrition Service, Office of Analysis and Evaluation, Evaluation of National Commodity Processing, June 3, 1983.

Walter Williams, et.al., Studying Implementation: Methodological and Administrative Issues, Chatham House Publishers, Inc., Chatham, NJ, 1982.